

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **07056599 A**(43) Date of publication of application: **03.03.95**

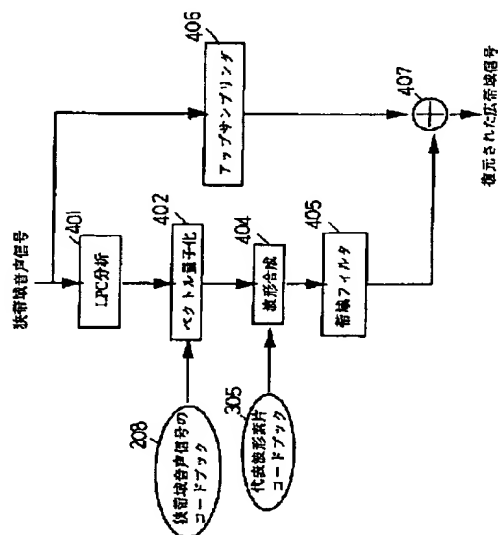
(51) Int. Cl.

G10L 9/14**G10L 9/18**(21) Application number: **05203199**(22) Date of filing: **17.08.93**(71) Applicant: **NIPPON TELEGR & TELEPH
CORP <NTT>**(72) Inventor: **ABE MASANOBU
YOSHIDA YUKI****(54) WIDE BAND VOICE SIGNAL RECONSTRUCTION
METHOD****(57) Abstract:**

PURPOSE: To reconstruct wide band voice signals (for example 8kHz/band) from narrow band voice signals (for example 300Hz to 3.4kHz) with a good quality.

CONSTITUTION: Input narrow band voice signals are LPC analyzed and their spectrum information parameters are obtained (401). These parameters are vector quantized using a narrow band voice signals code book 208 (402). Original voice waveforms corresponding to the code vectors are stored in a representative waveform piece code book 305 equivalent to one pitch for a voiced sound and equivalent to one frame for a voiceless sound for every code number of the code book 208. Using this code book 305 and output code vector numbers of the step 402, voiced sound are waveform synthesized by superimposing pitch synchronization and voiceless sounds are waveform synthesized by randomly using waveforms equivalent to one frame and wide band voice signals are obtained (404). Wide band components other than 300Hz to 3.4kHz are taken out from the wide band voice signals (405), synthesized with the amplified and sampled narrow band signals and wide band voice signals are reconstructed.

COPYRIGHT: (C)1995,JPO



(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平7-56599

(43)公開日 平成7年(1995)3月3日

(51)Int.Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 9/14	G	8946-5H		
	J	8946-5H		
9/18	E	8946-5H		

審査請求 未請求 請求項の数4 O L (全 9 頁)

(21)出願番号 特願平5-203199

(22)出願日 平成5年(1993)8月17日

(71)出願人 000004226

日本電信電話株式会社
東京都千代田区内幸町一丁目1番6号

(72)発明者 阿部 匡伸

東京都千代田区内幸町1丁目1番6号 日
本電信電話株式会社内

(72)発明者 吉田 由紀

東京都千代田区内幸町1丁目1番6号 日
本電信電話株式会社内

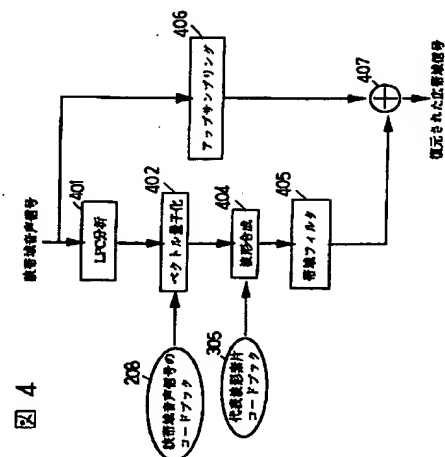
(74)代理人 弁理士 草野 卓

(54)【発明の名称】 広帯域音声信号復元方法

(57)【要約】

【目的】 狭帯域音声信号 (例えば300Hz~3.4kHz) から広帯域音声信号 (例えば8kHz帯) を高品質に復元する。

【構成】 入力狭帯域音声信号をLPC分析してスペクトル情報パラメータを得る(401)、そのパラメータを、狭帯域音声信号コードブック208を用いてベクトル量子化する(402)。コードブック208の各コード番号ごとに、そのコードベクトルと対応する原音声波形を、有声音は1ピッチ分、無声音は1フレーム分を代表波形素片コードブック305に記憶しておき、このコードブック305を用いて、ステップ402の出力コードベクトル番号について、有声音はピッチ同期の重ね合わせで、無声音は1フレーム分の波形をランダムに使用して波形合成して広帯域音声信号を得る(404)。その広帯域音声信号から300Hz~3.4kHz以外の広帯域成分を取り出し(405)、狭帯域信号をアップサンプリングしたものと合成して広帯域音声信号を復元する。



【特許請求の範囲】

【請求項1】 入力された狭帯域音声信号から広帯域音声信号を生成して出力する広帯域音声復元方法において、

入力された狭帯域音声信号をスペクトル分析する第1のステップと、

その第1のステップで得た結果を、予め用意した狭帯域音声信号のコードブックを用いてベクトル量子化する第2のステップと、

その第2のステップで得た量子化値を、予め用意した代表波形素片のコードブックを用いて復元して広帯域音声信号を得る第3のステップと、

からなる広帯域音声信号復元方法。

【請求項2】 上記入力された狭帯域音声信号をアップサンプリングを行ってサンプリング値を算出する第4のステップと、

上記第3のステップで得た広帯域音声信号から上記入力狭帯域音声信号帯域外の広帯域部分を取り出す第5のステップと、

その第5のステップで得た広帯域部分と上記第4のステップで得たサンプリング値とを加えて広帯域音声信号を得る第6のステップと、

を備えることを特徴とする請求項1記載の広帯域音声信号復元方法。

【請求項3】 上記代表波形素片のコードブックは、学習用広帯域音声信号をスペクトル分析し、そのスペクトル分析の結果と、予め用意した広帯域音声信号のコードブックとを比較し、その各コードについてそのパラメータと、最も近いスペクトル分析結果と対応する学習用広帯域音声信号を、有声音なら1ピッチ分、無声音なら1フレーム分抽出し、そのコードの代表波形素片としたものであることを特徴とする請求項1又は2記載の広帯域音声信号復元方法。

【請求項4】 入力された狭帯域音声信号から広帯域音声信号を生成して出力する広帯域音声復元方法において、

入力された狭帯域音声信号をスペクトル分析する第1のステップと、

その第1のステップで得た結果を、予め用意した狭帯域音声信号のコードブックを用いてベクトル量子化する第2のステップと、

その第2のステップで得た量子化値を、予め用意した広帯域音声信号のコードブックを用いて復元する第3のステップと、

その復元された広帯域音声信号から上記入力狭帯域音声信号より低域の成分を取出す第4のステップと、

上記第2のステップで得た量子化値を、予め用意した、狭帯域音声信号より高域の音声信号の代表波形素片のコードブックを用いて復元して高域音声信号を得る第5のステップと、

上記入力された狭帯域音声信号をアップサンプリングを行ってサンプリング値を算出する第6のステップと、

上記第4のステップで取出した低域成分と、上記第5のステップで復元した高域音声信号と、上記第6のステップで算出したサンプリング値とを加えて広帯域音声信号を得る第7のステップと、

を備えることを特徴とする広帯域音声信号復元方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 この発明は狭帯域音声信号から広帯域音声信号を生成する方法に関し、具体的には、現在電話音声やAMラジオ等で出力されているような狭帯域音声信号を、オーディオセットやFMラジオ等で出力されているような広帯域音声信号に高品質化することを可能とする方法に関する。

【0002】

【従来の技術】 狭帯域音声信号の例として電話音声について説明する。既存の電話システムが伝送できる信号のスペクトル帯域は、約300Hzから3.4kHzである。従来の音声の符号化技術の目的は、この電話帯域の音声の品質を保ち、かつ伝送パラメータ量を最小にすることであった。すなわち従来の音声の符号化技術では入力音声を再現することは可能であるが、入力音声の品質を超える音声を得ることは不可能である。一方、最近の音響技術の発展やディジタル処理の開発により日常生活で使われる音の品質が向上してきており、現状の電話帯域の音声の音質では満足できない状況が発生している。この要望を解決する方法としては、既存の電話システムを破棄し、広帯域の信号を伝送できるような電話システムを再構築することが考えられるが、経済的に大きな負担であるばかりでなく、再構築するにしてもかなりの時間を要すると考えられる。

【0003】

【発明が解決しようとする課題】 この発明の主たる目的は、例えば既存の電話システムを有効に利用して伝送された狭帯域音声信号を広帯域の音声信号として出力できるようにすること、また例えば広帯域の信号を伝送できるような電話システムと既存の狭帯域の電話システムとが共存する様な状況においても、両方の電話システムの組み合わせに関係なく、広帯域の音声信号を利用できるようにする広帯域音声信号復元方法を提供することにある。

【0004】

【課題を解決するための手段】 請求項1の発明によれば、第1のステップで入力された狭帯域音声信号をスペクトル分析し、その分析結果を、第2のステップで、予め用意した狭帯域音声信号のコードブックを用いてベクトル量子化し、その量子化値を、第3のステップで、予め用意した代表波形素片のコードブックを用いて復元して広帯域音声信号を得る。

【0005】請求項2の発明によれば、請求項1の発明において、更に第4のステップで、入力された狭帯域音声信号をアップサンプリングしてサンプリング値を算出し、第5のステップで、第3のステップで得た広帯域音声信号から入力狭帯域音声信号の帯域外の広帯域部分を取り出し、その取り出した広帯域部分を第6のステップで上記サンプリング値に加えて広帯域音声信号を得る。

【0006】請求項3の発明によれば、請求項1又は2の発明で用いる代表波形素片のコードブックを次のようにして作る。即ち学習用広帯域音声信号をスペクトル分析し、そのスペクトル分析の結果と、予め用意した広帯域音声信号のコードブックとを比較し、その各コードについてそのパラメータと最も近いスペクトル分析結果と対応する学習用広帯域音声信号を、有声音なら1ピッチ分、無声音なら1フレーム分抽出し、そのコードの代表波形素片とする。

【0007】請求項4の発明によれば入力狭帯域音声信号を第1のステップでスペクトル分析し、その分析結果を、第2のステップで、予め用意した狭帯域音声信号のコードブックを用いてベクトル量子化し、その量子化値を、第3のステップで、予め用意した広帯域音声信号のコードブックを用いて復元して広帯域音声信号を得、その広帯域音声信号から入力狭帯域音声信号より低域の成分を第4のステップで取出して低域信号を復元し、第2のステップで得た量子化値を、予め用意した狭帯域音声信号よりも高域の音声信号の代表波形素片のコードブックを用いて、第5のステップで高域信号を復元し、入力狭帯域音声信号を第6のステップでアップサンプリングを行ってサンプリング値を算出し、そのサンプリング値と、上記復元低域信号と、上記復元高域信号とを第7の* 30

$$D = \sum [C(i) - C'(i)]$$

ここで Σ は $i = 1$ から p まで、 C 及び C' は異なる音声信号をLPC分析して求めた各LPCケプストラム係数、 p はLPCケプストラム係数の次数である。なお、上述のLBGアルゴリズムについては、Linde, Buzo, Gray: "An algorithm for Vector Quantization Design" IEEE COM-28(1980-01)に詳細に記載されている。

【0011】上述の(1)式に基づいて、広帯域音声信号コードブック104が求まる。次に図2を参照して、広帯域音声信号コードブックとの対応関係ととりながら、狭帯域音声信号コードブックを作成する手順について説明する。まずステップ201において、学習用の広帯域音声信号から入力となる狭帯域音声信号と同一帯域の信号を作成する。この実施例においては広帯域音声信号を8kHz帯域の音声信号とし、狭帯域音声信号を電話帯域の音声信号として説明する。従って、ステップ201は300Hz以下の周波数を除去するハイパスフィルタと3.4kHz以上の周波数を除去するローパスフィルタとに広帯域音声信号を通すことによって実現される。一方、入力広帯域音声信号はステップ202においてLP 50

*ステップで加えて広帯域音声信号を得る。

【0008】

【実施例】この発明の方法では狭帯域音声信号のコードブックと、代表波形素片のコードブックとを必要とし、その代表波形素片のコードブックを作る際に広帯域音声信号のコードブックを用いる。従ってまずこれらコードブックの作成方法を説明する。

【0009】まず図1を参照して広帯域音声信号のコードブック作成手順について説明する。この作成手順は従来より知られ、広帯域音声信号の特徴を効率良く表現するために、広帯域音声信号の特徴を適切に表現するパラメータを用いてクラスタリングを行いコードブックを作成する。音声信号を特徴付けるパラメータとして線形予測分析(LPC)による音声スペクトル包絡や、FFTケプストラム分析法による音声スペクトル包絡、PSE音声分析合成法、正弦波の重ね合わせによる音声の表現法等が考えられるが、この実施例においては、LPCによる音声スペクトル包絡を特徴パラメータとして用いた場合について説明する。まず入力された広帯域、例えば8kHz帯域の音声はステップ101においてA/D変換器によってデジタル信号に変換される。その後、ステップ102においてLPC分析が施され、スペクトル情報(自己相関係数、LPCケプストラム係数)のパラメータが得られる。これらのパラメータを充分多く、例えば200単語程度について収集した後にステップ103においてクラスタリングを行う。クラスタリングはLBGアルゴリズムで行われるが、この際使用される距離尺度は(1)式で示すごとくLPCケプストラムのユークリッド距離 D である。

【0010】

$$\dots\dots (1)$$

C分析が施され、ステップ203において、前述の図1に示したコードブックの作成手順に従って求めた広帯域音声信号のコードブック104を用いて、ベクトル量子化される。

【0012】ところで、狭帯域音声信号は広帯域音声信号から作成されたものであるから、狭帯域音声信号と広帯域音声信号との時間対応はLPC分析を施すフレーム番号で1対1に対応をとることができる。この原理に従って、ステップ203でベクトル量子化した広帯域音声信号に対応する狭帯域音声信号を求め、この信号をステップ205でLPC分析し、その分析結果をステップ206において、ステップ203のベクトル量子化で得られたコードベクトル番号ごとに分類し保存する。つまり広帯域音声信号と狭帯域音声信号との時間対応とステップ202、205の両フレームとの対応と一致させ、同一フレーム番号の広帯域音声信号のベクトル量子化されたコードベクトル番号と、狭帯域音声信号のLPC分析結果とをそれぞれ対応させて保存する。以上、ステップ201からステップ206の処理を学習用に準備された

全ての広帯域音声信号、例えば200単語分に対して施す。ステップ207では、以上の全ての処理を通じてステップ206で保存されたLPC分析結果を、各クラス（同一コードベクトル番号）ごとに平均化処理を行い、その平均値をコードベクトルとして持つ狭帯域音声信号のコードブック208を作成する。

【0013】次に代表波形素片のコードブックを作成する手順を図3を参照して説明する。図1の広帯域音声信号のコードブックの作成に用いた学習用広帯域音声信号、又は異なるほぼ同一の周波数帯の学習用広帯域音声信号を入力してステップ301でA/D変換器によりデジタル信号に変換し、ステップ302でそのデジタル信号に付してLPC分析を行ってスペクトル情報（自己相関係数、LPCケプストラム係数）のパラメータを得る。このパラメータは図1のコードブックを作成した時のパラメータと同一のものとする。従って、図3のステップ301、302は省略し、図1のステップ102で得られたパラメータを用いてもよい。このパラメータを充分多く収集した後、例えば200単語程度について収集した後、ステップ303で、図1で作成した広帯域音声信号のコードブック104を参照して、各セントロイドに最も近いフレームを選ぶ、つまり、コードブック104の各コード番号について、そのコードベクトルに最も近いスペクトル情報パラメータを選択し、そのパラメータと対応した原広帯域音声信号における1分析窓長、（LPC分析における）の波形を、有声音ならば1ピッチ分、無声音ならば1分析窓長分抽出し、その抽出波形をそのコード番号に対する代表波形素片とする。このようにして各コード番号に対する代表波形素片を収容した代表波形素片コードブック305を得る。

【0014】次に請求項1の発明により狭帯域音声信号から広帯域音声信号を復元する手順を図4を参照して説明する。入力された、例えば300Hz～3.4kHzの狭帯域音声信号を、ステップ401においてLPC分析して図1で用いたものと同様のスペクトル情報を得、この得られたスペクトル情報パラメータを、ステップ401においてベクトル量子化する。このベクトル量子化では、図2で示した方法で作成した狭帯域音声信号コードブック208が使用される。次に、ステップ404において、図3で求めた代表波形素片305を用いてコードベクトル番号と対応する代表波形素片を取り出し、有声音部分はピッチ同期の重ね合わせで、無声音部分は窓長シフト幅（LPC分析における）分の波形をランダムに使用することによって波形合成して例えば8kHz帯域の広帯域音声信号を得る。

【0015】以上の処理で求めた広帯域音声信号は、狭帯域音声信号には存在しない信号を含んでいるが、狭帯域音声信号に存在していた信号を歪ませるという副作用を起こす。そこで、請求項2の発明では次に述べる処理を行う。すなわちステップ405において、ステップ

404で得られた広帯域音声信号を、300Hz以下の周波数を取り出すローパスフィルタと3.4kHz以上の周波数を取り出すハイパスフィルタとに通し、入力狭帯域音声信号の帯域外の広帯域部分を取り出す。一方、入力狭帯域音声信号をステップ406で8kHz帯域にアップサンプリングし、そのサンプリング値と、ステップ405の出力広帯域部分とをステップ407でたしあわせて、復元された広帯域音声信号を得る。

【0016】上述において、代表波形素片コードブックによる入力信号帯域外の復元を高域側のみとし、低域側は、音声素片を用いなくてもよい。即ち例えば図5に示すように図4について述べたように、入力狭帯域音声信号をLPC分析し、（401）その分析結果を狭帯域音声信号コードブック208を用いてベクトル量子化するが（402）、その量子化コードを、広帯域音声信号コードブック104を用いて、復元し、つまり、この量子化コードと同一コードの広帯域音声信号コードブック104のコードベクトルでLPC音声合成器のフィルタ係数を制御し、そのLPC音声合成器にステップ401でのLPC分析にもとづくピッチの励振信号を入力し、また出力レベルをステップ401のLPC分析のレベルに応じて制御する（501）。このように復元された広帯域音声信号から狭帯域入力信号より低域成分、例えば300Hz以下を低域通過フィルタで取出す（502）。この時、ステップ401でのLPC分析による検出パワーは300Hz～3.4kHzの入力狭帯域信号のパワーであり、この、パワーとステップ501からの、LPC合成信号、例えば8kHz帯の信号がパワーが等しくなるようにされる。よってこの復元された広帯域音声信号のパワーレベルは、入力狭帯域信号の帯域についてみると低下したものとなっている。その低下した分をステップ502の低域通過フィルタ出力に対して調整して入力狭帯域音声信号のパワーレベルと対応するように上げる（503）。このようにして入力狭帯域音声信号と対応したその帯域よりも低い帯域の信号が復元される。

【0017】次に、入力狭帯域音声信号と対応するその帯域よりも高域の信号を復元するために二つの代表波形素片コードブックを予め作成しておく。つまり、図3に示すように、学習広帯域音声信号から代表波形素片コードブック305を作る場合と、同様の手法で図6に示すように、学習広帯域音声信号を、広帯域音声信号コードブック104を用いてベクトル量子化し、その各コードについてコードベクトルを最も近い学習広帯域音声信号の波形を有声音については1ピッチ分、無声音についてはLPC分析における1分析窓長分抽出する（303）。この抽出した各代表波形素片を、例えば300Hz～3.4kHzを通過帯域とするフィルタに通して（601）、狭帯域の代表波形素片コードブック602を作り、またステップ303で抽出した各代表波形素片を例えば3.4kHz以上を通過させる高域通過フィルタに通

して(603)、高域の代表波形素片コードブック604を作る。

【0018】これら兩代表波形素片コードブック602、604を用いて、入力狭帯域音声信号から、その帯域よりも高い帯域の信号を復元する手法を図7を用いて説明する。図4のステップ402で入力狭帯域信号を、狭帯域音声信号コードブック208を用いてベクトル量子化した、その量子化コード番号を用いて、狭帯域代表波形素片コードブック602から代表波形素片を選択し(701)、また前記量子化コード番号を用いて、広帯域の代表波形素片コードブック604から代表波形素片を選択する。これら選択した狭帯域、広帯域の各代表波形素片がそれぞれ無声音か有声音かの判定を行う(703、704)。無声音と判定されると、それぞれその選択した狭帯域、広帯域の各代表波形素片から、窓シフト幅分ずつ、開始点をランダムに選定して取出す(705、706)。ステップ703、704で有声音と判定されると、その選択した狭帯域、広帯域の各代表波形素片を、LPC分析(図4のステップ401)の結果得られているピッチと同期して取出して重ね合わせる(707、708)。ステップ705、707で取出した各代表波形素片列のパワーと、LPC分析(図4のステップ401)の結果得られているパワーとの比をそれぞれ求める(709、710)。これら求めた比を、ステップ706、708より得られ、代表波形素片列のレベルに掛算して、入力狭帯域音声信号のパワーと対応するパワーとして(711、712)、加算することにより入力狭帯域音声信号と対応するその帯域の高域側の復元信号を得る(713)。

【0019】この高域側復元信号と、図5で得られた低域側復元信号と、図4中のステップ406で得られた狭帯域音声信号とを加えて広帯域音声信号を得る。図1中のステップ102、図2中のステップ202、205、図3中のステップ302、図4中のステップ401における各スペクトル分析は同一分析法により同種のパラメータを求める。図2の狭帯域音声信号コードブック208の作成に用いる学習用広帯域音声信号は、広帯域音声信号コードブック104の作成に用いた広帯域音声信号を用いることが好ましい。何れにしても両音声信号の特徴の対応関係を保存しながら両コードブックを作成するとよい。しかし、この場合より音質が多少悪くなるが、*

(1)	(2)
6.9%	96.9%

(3)の本発明による復元音声、(4)のLPC合成音による復元音声は7.3kHz帯域音声に近いという結果が、それぞれ、86.2%、86.4%であり、復元音声は、7.3kHz帯域音声に近いという結果が86.4%であり、この発明によれば、可成りよく、帯域の復元がなされていることが理解される。

【0023】以上述べたように、この発明によれば、狭

*広帯域音声信号のコードブック104と、狭帯域音声信号のコードブック208の各作成に全く別の音声信号を用いてもよく、かつ狭帯域音声信号のコードブック208を図2に示したように、広帯域音声信号と狭帯域音声信号の特徴の対応関係を保存させて作成するのではなく、図1に示した通常の手法で狭帯域音声信号コードブック208を作ってもよい。このようにしても広帯域音声信号と狭帯域音声信号とは、例えば同一音韻についてみればその特徴は一般的にかなり相関があり、狭帯域音声信号の同一音韻について広帯域音声信号のコードブック中の同一音韻を用いれば音質がかなり向上することが期待できる。この場合広帯域音声信号のコードブック104と狭帯域音声信号のコードブック208との各同一コード番号について両コードベクトルを対応付けておく必要がある。

【0020】広帯域音声コードブック104を作るための広帯域音声信号と、代表波形素片コードブック305を作るための広帯域音声信号とは同様に同一でも異なってもよい。

【0021】

【発明の効果】学習データとして音韻バランス単語186単語を用い、分析窓にハミング窓を用い、分析窓長を21ミリ秒とし、窓シフト幅を3ミリ秒とし、LPC分析次数を14次とし、FFTポイント数を512とし、コードブック作成時の距離尺度をLPCケプストラムユークリッド距離とし、広帯域音声信号コードブック104のサイズを16、狭帯域音声信号コードブック206のサイズを256として、次の試験を行った。

【0022】A及びBとして電話帯域音声、7.3kHz帯域音声ランダムに提示する。3番目に提示される音声Xを次の(1)から(4)とした。

- (1) 電話帯域音声
 - (2) 7.3kHz帯域音声
 - (3) 図5、図7の復元法(本発明)を利用した音声
 - (4) (2)のLPC分析合成音の低域、高域を電話帯域音声に加え合わせた音声
- (4)がLPC系を利用した場合に最良の復元音声になると考え、被験者にはXの音声はAに近いのかBに近いのかを判断させた。全部で125組の音声をヘッドホンを用いて提示した。被験者は6名である。Xが7.3kHz帯域音声に近いと判断した割合を下記に示す。

(3)	(4)
86.2%	86.4%

帯域音声信号と、広帯域音声との特徴の対応により、狭帯域音声信号に存在しない音声信号の特徴を効率よく復元することができる。特に、広帯域音声信号をそのまま音声波形素片として用いて復元するため高品質な復元音声を得ることができる。

【図面の簡単な説明】

【図1】広帯域音声信号コードブックを作成する手順を

示す図。

【図2】狭帯域音声信号コードブックを作成する手順を示す図。

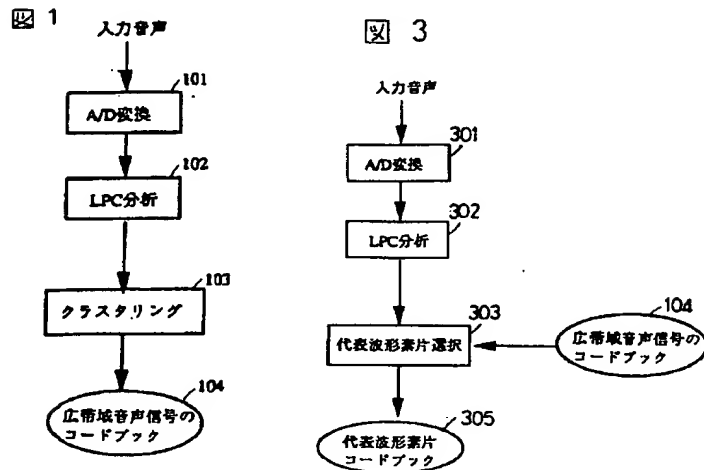
【図3】代表波形素片コードブックを作成する手順を示す図。

【図4】請求項1及び2の各発明による広帯域音声信号復元方法の手順を示す図。

【図1】

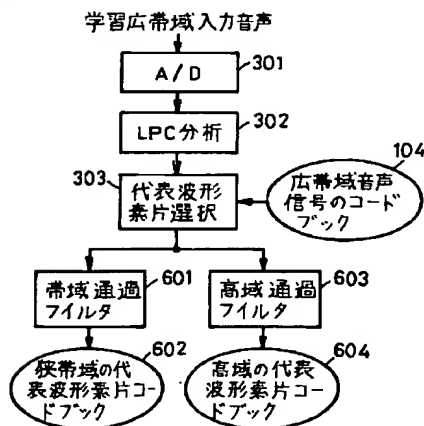
【図3】

【図5】



【図6】

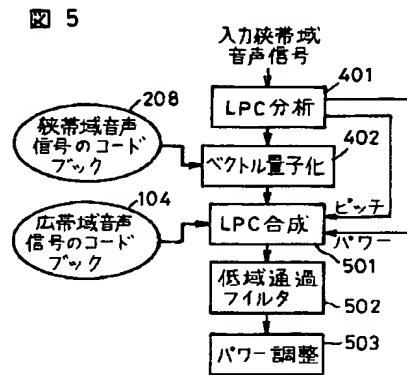
図 6



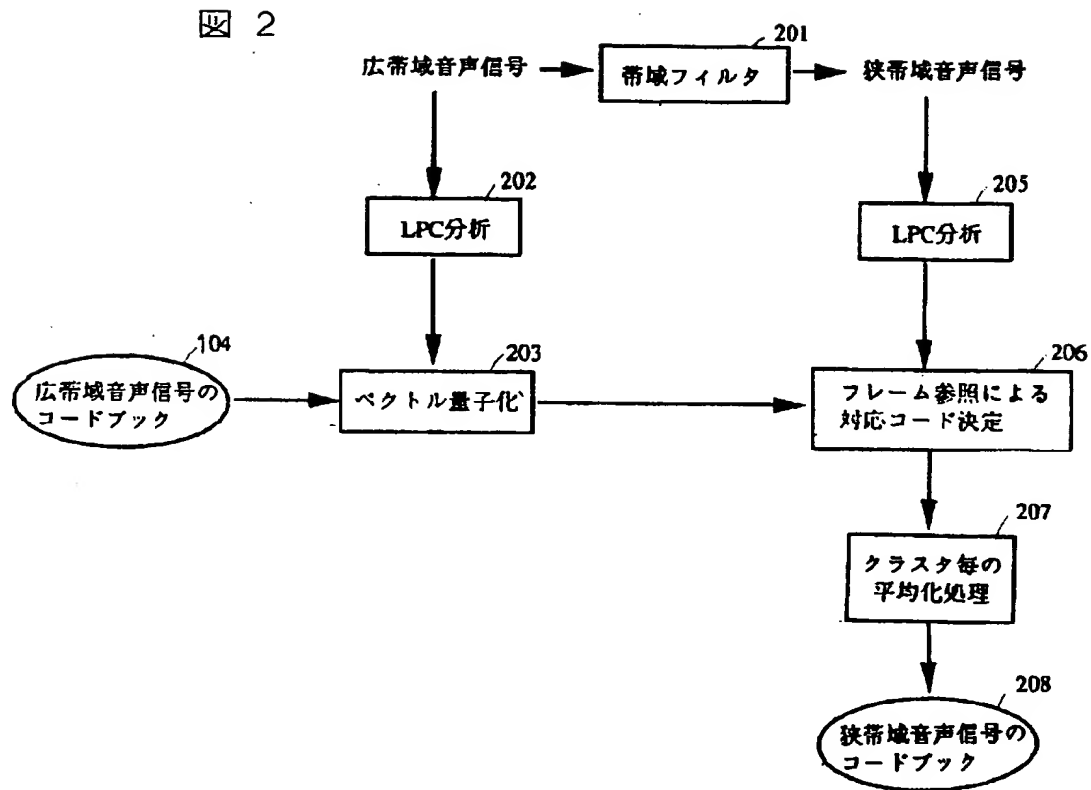
【図5】入力狭帯域音声信号よりも低い帯域の信号を復元する方法の手順を示す図。

【図6】狭帯域の代表波形素片コードブック及び高域の代表波形素片コードブックを作成する手順を示す図。

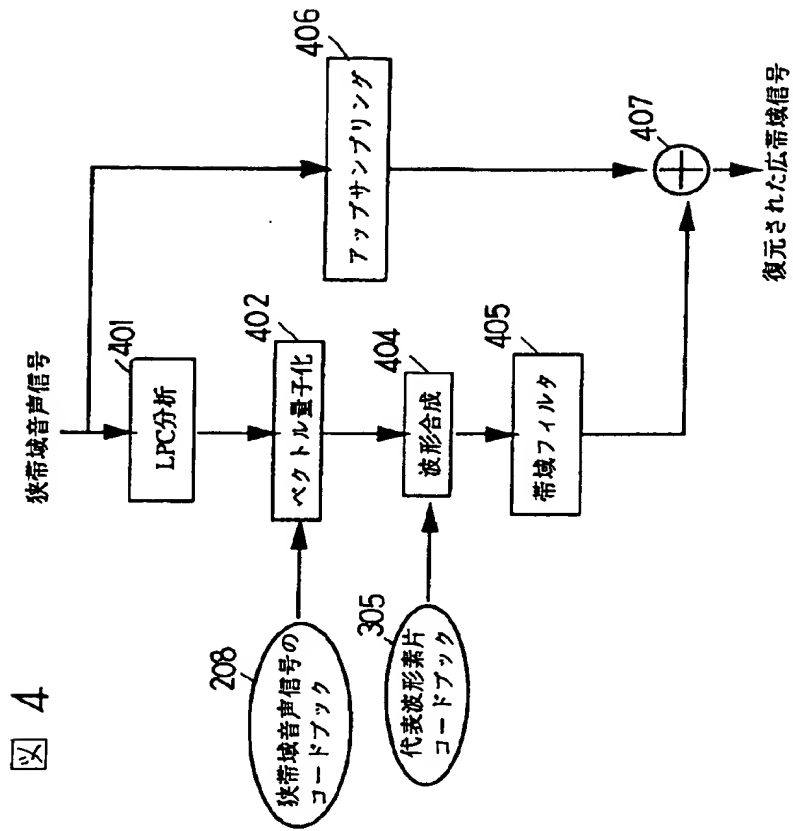
【図7】入力狭帯域音声信号よりも高い帯域の信号を復元する方法の手順を示す図。



【図2】



【図4】



【図7】

